Chapter 21
Vagueness, Truth and Permissive Consequence

Pablo Cobreros, Paul Egré, David Ripley and Robert van Rooij

Abstract We say that a sentence A is a *permissive consequence* of a set of premises Γ whenever, if all the premises of Γ hold up to some standard, then A holds to some weaker standard. In this paper, we focus on a three-valued version of this notion, which we call *strict-to-tolerant* consequence, and discuss its fruitfulness toward a unified treatment of the paradoxes of vagueness and self-referential truth. For vagueness, st-consequence supports the principle of tolerance; for truth, it supports the requisit of transparency. Permissive consequence is non-transitive, however, but this feature is argued to be an essential component to the understanding of paradoxical reasoning in cases involving vagueness or self-reference.

21.1 Introduction

According to the standard view of logical consequence, a sentence A is said to follow from a set of premises Γ if it is impossible for all the premises of Γ to be true together and for the conclusion A *not to be true*. Alternatively, a sentence A may be said to follow from a set of premises Γ if it is impossible for all of the premises of Γ to be true together and for the conclusion A to be *false*. In a bivalent setting, the foregoing definitions coincide, because falsity and non-truth coincide. When the underlying space of truth values is larger, however, these two definitions can come apart, and the

P. Egré
Institut Jean Nicod, Ecole Normale Supérieure-PSL Research University, Paris, France
e-mail: paul.egre@ens.fr

P. Cobreros
Universidad de Navarra, Pamplona, Spain
e-mail: pcobreros@unav.es

D. Ripley
University of Connecticut, Storrs, CT, USA
e-mail: davewripley@gmail.com

R. van Rooij
University of Amsterdam, Amsterdam, Netherlands
e-mail: R.A.M.vanRooij@uva.nl

T. Achourioti et al. (Eds.), *Unifying the Philosophy of Truth,* Logic, Epistemology, and the Unity of Science 36, DOI 10.1007/978-94-017-9673-6_21

latter definition, in particular, becomes potentially more permissive than the former, allowing for more schemata to count as valid inference patterns.

Our goal in this paper is to show the interest of such a notion of permissive consequence, whereby consequence is no longer defined as the preservation of some designated truth-value (or set thereof) from premises to conclusion, but rather, as the enlargement of the set of designated truth-values, or as a weakening of standards when going from premises to conclusion (see, in order of appearance, (Nait-Abdallah 1995), (Bennett 1998), (Frankowski 2004), (Zardini 2008), (Smith 2008), (van Rooij 2012), (Cobreros et al. 2012b)). More specifically, we intend to show the fruitfulness of this notion for the prospect of getting a unified treatment of the paradoxes of vagueness and of the paradoxes of self-referential truth. The notion of permissive consequence we are concerned with was originally introduced with an aim to solving the sorites paradox (see (Zardini 2008)), and in order to account for the semantics and pragmatics of vague predicates more generally (see (vanRooij 2012), (Cobreros et al. 2012b), (Cobreros et al. 2012a)). It soon became apparent, however, that it could be applied in a natural way to the treatment of the semantic paradoxes and in particular to the Liar Paradox (see (Ripley 2012), (Cobreros et al. 2013)).

The quest for a unified treatment of vagueness and self-referential truth has been viewed as both natural and desirable by several authors before us (see in particular (McGee 1991), (Tappenden 1993), (Field 2003), (Colyvan 2009), (Priest 2010)). One of the reasons for that is that both the paradoxes of vagueness and the semantic paradoxes appear to put into question two central laws of classical logic, namely the law of excluded middle and the principle of non-contradiction. In the case of vagueness, borderline cases often appear as semantically indeterminate cases, cases of which it is neither determinately true to say that the predicate holds, nor determinately false to assert it. The same appears to hold of our use of the word 'true'. As McGee put it, there are sentences "that the rules of our language, together with the empirical facts, determine to be definitely true; sentences that the rules of our language, together with the empirical facts, determine to be definitely not true; and sentences that are left unsettled" (McGee 1991, p. 6). Among those, Liars and Truth-Tellers figure most prominently.

For vagueness as well as for truth, consequently, three-valued logic appears as a natural and well-motivated framework. The addition of a third truth value to deal with vague predicates or with the truth predicate leaves a number of issues open, however, starting with the interpretation of the third truth value, and with the choice of an appropriate consequence relation. Let us agree to call the value 1 'true-only', value 0 'false-only', and leave open what to call the value $\frac{1}{2}$ (see (Priest 1979), (Lewis 1982)). Depending on the view, $\frac{1}{2}$ may be called 'neither true nor false', or 'both true and false'. On paracomplete approaches, the Liar sentence is fundamentally viewed as neither true nor false, and borderline cases of vague predicates are cases for which it is neither true nor false that the predicate applies. On the dual, paraconsistent approaches to vagueness and the Liar paradox (Priest 1979), the Liar sentence is fundamentally viewed as both true and false, and similarly borderline cases of vague predicates are cases for which it is both true and false that the predicate applies. Importantly, distinct logics result depending on which interpretation of the third truth value is favored, and on whether logical consequence is defined as the preservation of the value 1 (the true-only), or as the preservation of non-0 values (the non-(false-only)).

As it turns out, however, the duality between paracomplete and paraconsistent logics is such that the relative merits that one logic may claim over the other can usually be turned into relative limitations, and conversely. We will argue that a promising avenue for the treatment of the paradoxes lies in the definition of a consequence relation that results from a combination of paracomplete and paraconsistent features, rather than in the choice of one approach exclusive of the other. The relation of permissive consequence we have in mind is exactly along those lines, since it requires that when the premises of an argument take value 1 (are true-only), the conclusion must not take value 0 (is not false-only). A striking feature of this permissive consequence relation is that it exactly coincides with classical logic when no special provisos are made to deal with vagueness or with self-referential truth. When such provisos are included, however, we will see that this notion only departs from classical logic in that it yields a nontransitive consequence relation. This feature, arguably, does not constitute a cost: rather, we will argue that it captures a common and fundamental aspect to both families of paradoxes.

The paper is structured as follows. In Sect. 2, we give a brief overview of three-valued logic and introduce the notion of permissive consequence or st-consequence we use as our framework. In Sect. 3 we show how to extend the basic framework to accommodate vagueness on the one hand, and self-referential truth on the other, and in particular to deal with the Liar paradox and the sorites paradox. In Sect. 4, finally, we propose an assessment of our approach with regard to two main issues: the nontransitive character of permissive consequence on the one hand, and so-called revenge problems on the other, namely the treatment that we can give in our framework of the strengthened Liar and of higher-order vagueness.

21.2 Permissive Consequence and the Logic ST

21.2.1 The Scope of Permissive Consequence

The shape of the notion of permissive consequence we are about to introduce is by no means specific to the framework of three-valued logic. Also, in the literature the notion comes under various other names, such as *plausible* consequence (Frankowski 2004), *potential* consequence (Nait-Abdallah 1995), *arguable* consequence (Bennett 1998), or *tolerant* consequence (Zardini 2008). It can be defined for any logic in which it is sensible to distinguish a set of designated values and a set of tolerated values, where the set of tolerated values is a superset of the set of designated values. The general form of such a consequence relation was introduced and investigated independently by Frankowski (in (Frankowski 2004), based on earlier work by (Malinowski 1990) on the dual notion of *quasi*-consequence) and by Zardini in (Zardini 2008), with rather minimalist assumptions about the algebra of truth-values in each case. Zardini does not make restrictions, in particular, about the number of truth-values, nor on whether truth-values should be partially or linearly ordered. Different logics correspond to this notion of permissive consequence depending on the algebra under consideration. Three-valued logic, however, is in a sense the smallest non-trivial framework for the investigation of this notion of permissive consequence, and interestingly, this notion was introduced independently by (Nait-Abdallah 1995) and by (Bennett 1998) in a trivalent setting.[1]

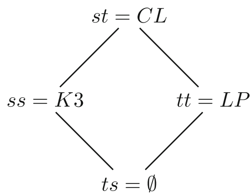
Bennett in particular put forward a notion of 'arguable entailment' for supervaluations, which he defined as follows: "if all of the premises are 'unequivocally' true, then the conclusion is 'in some sense' true". Although Bennett does not present it in that way, the definition can be seen to combine the notions of super-truth and sub-truth that are familiar from the literature on vagueness (see (Hyde 1997), (Cobreros et al. 2012a)). Even closer to our proposal, in an underappreciated book Nait-Abdallah investigated the interplay between two notions of truth in a trivalent setting, which he calls *classical truth* (for the value 1) and *potential truth* (for values > 0), which correspond exactly to the notions of strict truth and tolerant truth we are about to review and that we introduced independently. In different ways, however, Nait-Abdallah's study is both more general and more restricted than ours. It is more general in that it studies a notion of consequence in which premises and conclusions can be interpreted strictly or tolerantly in a non-uniform manner. It is more specific in that Nait-Abdallah limited his study of permissive consequence—consequence from classical to potential truth—to the case of consequence from zero premises.

21.2.2 *st*-consequence

To make our definitions precise, let us consider the language of first-order logic without identity and function symbols as our basic language, with negation, conjunction, and the universal quantifier as our basic logical connectives (we assume that the symbols $\bot$ and $\top$ are not part of the language). Importantly, we define the conditional $\supset$ as the material conditional in the usual way. We define three-valued models for this language over the set $\{1, \frac{1}{2}, 0\}$ of truth-values, using Kleene's strong schema as our valuation schema. According to Kleene's strong schema, negation maps the value 1 to 0, 0 to 1, and $\frac{1}{2}$ onto itself; conjunction is defined as the minimum of the values of the conjuncts, and universal quantification as the minimum of values over all assignments that differ at most on the value they assign to the variable bound by

[1] (Smith 2008), for instance, defines a notion of permissive consequence for fuzzy logic with continuum many truth values linearly ordered (the real interval $[0, 1]$), whereby a sentence A is said to follow from a set of premises Γ provided in all models in which all formulae in Γ have value strictly greater than half, A gets a value greater or equal than 1/2. This notion of permissive consequence, as fine-grained though it appears, can in fact be shown to be representable without loss of generality in a three-valued framework, and indeed, both Smith's notion and its three-valued version coincide with classical consequence for first-order logic. We refer to (Cobreros et al. Ms.) for details and more ample discussion of this point.

Fig. 21.1 Mixed consequence based on s and t



the quantifier (viz. (Kleene 1952)). A Kleene model for first-order logic is a structure $M = (D, I)$ where D is a set of individuals, and I an interpretation function for the non-logical vocabulary, that maps n-ary predicate symbols to functions from D^n to $\{1, \frac{1}{2}, 0\}$.

Based on this, let us say that a sentence A is *strictly true* or s-true in M, noted $M \models^s A$, provided $I(A) = 1$; we say that it is *tolerantly true* or t-true in M, noted $M \models^t A$, provided $I(A) > 0$. Thus, s-truth corresponds to what we earlier called for a sentence to be true-only, and t-truth for a sentence to be non-(false-only). Clearly, s-truth and t-truth are duals, that is, a sentence is tolerantly true iff its negation is not strictly true, and vice versa. For $n, m \in \{s, t\}$, moreover, the usual notion of logical consequence can be generalized by saying that:[2] $\Gamma \models^{nm} \Delta$ provided there is no model M such that $M \models^n \gamma$ for every $\gamma \in \Gamma$ and $M \models^m \delta$ for no $\delta \in \Delta$. As shown in Fig. 21.1, we thereby get four distinct notions of 'mixed' consequence.

When $n = m = s$, the resulting notion of logical consequence, or preservation of strict truth from premises to conclusions, coincides with Kleene's strong logic K3. When $n = m = t$, logical consequence corresponds to preservation of non-falsity or tolerant truth from premises to conclusions and the resulting system is Priest's Logic of Paradox (LP). When $nm = ts$, this corresponds to a case in which we go from tolerantly true premises to strictly true conclusions. The corresponding relation of consequence can be shown to be empty in this case. Intuitively, this corresponds to a notion of *restraining* consequence, since conclusions have to match a higher standard for truth than the premises. Conversely, the notion of *permissive* consequence we elect is defined in a dual way, namely as st-consequence, in that it asks for strictly true premises to imply conclusions that are tolerantly true.

The remarkable feature of st-consequence is that it coincides with classical consequence. Obviously, a classical countermodel to the entailment from Γ to Δ is an st-countermodel. But conversely, any st-countermodel can be turned into a classical countermodel, basically because reassignments of the values 1 or 0 to subsentences with value $\frac{1}{2}$ in the original model do not alter the value 1 or 0 assigned to the sentences in which they appear.

[2] We state the definition in terms of a multi-premise and multi-conclusion setting, although, for most of the applications we are interested in here, we can limit our perspective to the multi-premise-single-conclusion case.

Let us call ST the logic of st-consequence, and TS the logic of ts-consequence. It is worth stressing two aspects in which ST improves on the non-classical behavior of LP and K3. Both LP and K3 differ from CL in that both lose some classical validities, and both lose the deduction theorem. In particular, although $A \models_{K3} A$, it is not the case that $\models_{K3} A \supset A$. Similarly, although $\models_{LP} (A \wedge A \supset B) \supset B$, it is not the case that $A, A \supset B \models_{LP} B$. Thus, in K3 the loss of the deduction theorem is related to the loss of excluded middle in the same way in which, in LP, the loss of the deduction theorem is related to the loss of modus ponens as a valid inference. Because of that, it is generally agreed that neither K3 nor LP provides a satisfactory analysis of the conditional.

Upon reflection, this deficiency is not surprising. Indeed, it is easy to see that a conditional of the form $A \supset B$ is tolerantly true in a model provided either A is not strictly true, or B is tolerantly true, that is, provided that if A is strictly true, then B is tolerantly true. Dually, a conditional $A \supset B$ is strictly true in a model provided if A is tolerantly true, then B is strictly true. This suggests that in so far as a consequence relation can be expected to mirror the semantic behavior of its object-language conditional, st-consequence is the right correlate of the material conditional of LP, whereas ts-consequence is the right correlate for the material conditional of $K3$. In the case of ST, in particular, we will see in Sect. 3 that the classical behavior of the conditional is a significant advantage when we deal with vagueness and self-referential truth.

21.2.3 What Do 'strict' and 'tolerant' Mean?

A few more words are in order regarding the way in which the notions of strict truth and tolerant truth should be understood. We defined permissive consequence as the entailment from strict truth to tolerant truth. This may raise the legitimate worry that truth becomes an ambiguous notion.

Importantly, talk of tolerant *truth* and strict *truth* is not required to make sense of the notion of permissive consequence. An alternative route consists in linking the semantic values 1 and 0 not to truth proper, but to assertion. If we do so, "tolerant truth" and "strict truth" become essentially a *façon de parler*, and should be understood as shorthand for tolerant assertion and strict assertion. The idea, basically, is that a sentence is assertible strictly when there is non-arbitrary ground for the assertion. It can be denied strictly if there is non-arbitrary ground for denying it. To say that a sentence can be asserted or denied tolerantly means that there is ground for the assertion, but ground that may contain some element of arbitrariness (such as the existence of equal ground for the opposite assertion). Finally, a sentence can be such that it is assertible tolerantly and deniable tolerantly (at the opposite, no sentence can be asserted and denied strictly, but both a sentence and its negation can fail to be assertable strictly). Sentences that fall in that third category, we shall argue in the next section, are best matched by those sentences for which the rules that connect our use of language to empirical facts leave room for unsettedleness (remember the quote by McGee above). In what follows, we will see that we can use the strict/tolerant distinction as a way of classifying problematic sentences involving either vague predicates or self-referential truth.

The interpretation of the strict/tolerant distinction in terms of assertability rather than truth is compatible with an inferentialist interpretation of logical consequence, as opposed to what we might call a referentialist conception, on which truth values essentially reflect the correspondence status between a sentence and a state of affairs. By inferentialism, we mean the view on which linguistic meanings are to be explained by which inferences are valid, and more specifically the *bilateralist* view on which the validity of arguments itself is to be explained by general constraints on the speech acts of assertion and denial, or acceptance and rejection more generally (see (Rumfitt 2000), (Ripley 2013b), and (Malinowski 1990)).[3] This interpretation is arguably the most adequate when it comes to incorporating a transparent truth predicate in the language (see (Ripley 2012) and (Cobreros et al. 2013) and below for more ample discussion). This interpretation is not mandated, however. Some may find more appeal in the distinction between strict truth and tolerant truth as two levels of truth proper. (Smith 2008) for example defends a notion of permissive consequence for fuzzy logic in writing (p. 223):

> a sentence needs to meet more stringent standards of truth if it is to be used as the basis for further argument than if it is merely to be asserted—just as building codes place more stringent standards of load-bearing capacity on foundations than on superstructures.

Given Smith's commitment to degrees of truth, by standards of truth we take Smith to mean that assertability is based on those different levels of truth proper. A more neutral conception is defended by (Zardini 2008) who prefers to talk of truth values as "levels of goodness", whereby goodness is essentially a measure of the normative attitude to take toward a sentence (whether to believe it, assert it, or act upon it, see p. 345), without those attitudes necessarily being called good by reference to the truth of the corresponding sentence. If such levels of goodness are seen as ways of linking assertion to grounds for assertion, then they readily fit a unitary conception of truth, but a dual conception of assertion, on which the latter can come with different force.

21.3 Vagueness and Truth

Over a three-valued architecture, we see that ST allows us to preserve a classical notion of logical consequence. In this section we show how to extend ST to deal specifically with vague predicates on the one hand, and with a truth predicate on the other. In the case of vagueness, we will see that ST allows us to accommodate the

[3] Malinowski's notion of q-consequence is defined exactly in terms of the basic attitudes of acceptance and rejection. A sentence A is a q-consequence of a set of premises Γ whenever A is accepted when all of the premises in Γ are not rejected.

tolerance principle. In the case of truth, it allows us to accommodate a transparent truth predicate. In this section, we briefly review how to extend ST to deal with either kind of predicate. We mostly stress the analogies. We postpone a discussion of the potential disanalogies and limits of our account until the next section.

21.3.1 STVP

The hallmark of vague predicates on our account is the *tolerance principle* (Wright 1976), according to which a sufficiently small shift of the P-relevant respects of an individual should not make a difference as to whether the predicate P can be applied to that individual. For example, if someone 178cm tall is to be considered "tall", then someone only slightly shorter (177cm) can be considered tall too. More generally, we take the principle to be that if P is applicable to x, and x is sufficiently similar to y in the relevant respects, then P is applicable to y as well. In two-valued classical logic, the tolerance principle leads to paradox. In our approach, the tolerance principle can be validated without paradox.

To see this, we proceed to define an extension of ST called STVP (for 'ST with Vague Predicates'). As our language, we consider the language of first-order logic without identity, and containing, for simplicity, only unary predicates.[4] For each unary predicate P, moreover, the language contains a similarity predicate I_P. The formula aI_Pb is to be interpreted as: a and b are indiscriminable or sufficiently similar in P-relevant respects. Given a three-valued model $M = (D, I)$, each such predicate is to be interpreted by a relation $\sim_P$, with the following proviso:

$$M \models^s aI_Pb \text{ iff } M \models^t aI_Pb \text{ iff } |I(Pa) - I(Pb)| < 1 \text{ (closeness)} \tag{21.1}$$

Three comments can be made on this definition of P-similarity. First, it implies that two individuals are P-similar on that view provided the application of the predicate P yields truth values that are sufficiently close. This interpretation of P-similarity in terms of closeness in truth values is faithful to what Smith calls the *closeness principle*, according to which, if two individuals a and b are sufficiently similar in P-relevant respects, then the degrees of truth of the corresponding sentences Pa and Pb should not be too far apart (Smith 2008). Secondly, the relation of similarity is, for each predicate, reflexive and symmetric, but it need not be transitive. The non-transitivity of indiscriminability or similarity is actually a central aspect to our conception of vagueness, and we share it with significantly different accounts of the logic of vague predicates (see (Williamson 1994) in particular). Thirdly, this relation of P-similarity has a crisp interpretation, in the sense that it makes no difference whether it is interpreted strictly or tolerantly.[5]

---

[4] The generalization to n-ary predicates presents no special difficulty, and only involves the definition of appropriate similarity relations between n-tuples.

[5] For more rigor, we might have chosen to break the proviso 21.1 into two separate constraints: first, the crispness constraint that $M \models^s aI_Pb$ iff $M \models^t aI_Pb$ (see (Cobreros et al. 2012b)), and

Let us call $\models_{STVP}$ the relation of ST-consequence specific to this language, with the proviso 21.1. It is easy to see that:

$$Pa, aI_Pb \models_{STVP} Pb \text{ (tolerance as a rule)} \tag{21.2}$$

and similarly that:

$$\models_{STVP} \forall xy(Px \land xI_Py \supset Py) \text{ (tolerance as an axiom)} \tag{21.3}$$

For this means that if Pa gets value 1 in the model, and the distance in truth values between Pa and Pb is less than 1, then the truth value of Pb is necessarily greater than 0. Because st-valid formulae coincide with tt-valid formulae, note that the validity of the tolerance principle could have been obtained with a tt or tolerant-to-tolerant consequence relation, that is using LP as background logic. However, the same difference between ST augmented with vague predicates and LP augmented with vague predicates remains that we had between ST and LP: in ST for the language with vague predicates, the conditional satisfies modus ponens, and more generally it satisfies the deduction theorem. This feature matters particularly, for as eloquently argued by (Zardini 2008, p. 339), rejection of modus ponens in the case of vagueness seems to "deprive" the tolerance principle, formulated in conditional form, "of its intended force".

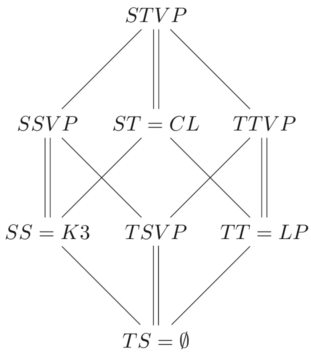
More generally, for each of the four logics we considered above, namely ST, TT, SS and TS, the inclusion of similarity predicates with the closeness proviso in 21.1 yields four new consequence relations, which we call STVP, TTVP, SSVP and TSVP. The latter strictly extend the former (see Fig. 21.2 below; for instance TSVP now has validities such as $aI_Pb \supset aI_Pb$), and moreover, these are (model-theoretic) conservative extensions, in the sense that they coincide on the set of I_P-free formulae (we mark this relation of conservative extension with double lines on Fig. 21.2).

Obviously, although ST coincides with CL, STVP no longer coincides with CL, in particular because the tolerance principle is not classically valid. A further central difference between STVP and CL is that STVP does not yield a transitive consequence relation. This explains, in particular, why the sorites paradox can be blocked. For instance, we have that $aI_Pb, Pa \models_{STVP} Pb$ and $bI_Pc, Pb \models_{STVP} Pc$ but $aI_Pb, bI_Pc, Pa \not\models_{STVP} Pc$ (assume that $I(Pa) = 1$, $I(Pb) = 1/2$, $I(Pc) = 0$).

The nontransitive feature of STVP is intuitively faithful to the nontransitive character of indiscriminability in this example. This means that although the tolerance principle is $STVP$-valid, the tolerance step cannot be taken more than once without risk when reasoning with vague predicates. Note that because STVP satisfies the deduction theorem, the two versions of tolerance stated above, 21.2 and 21.3, are equivalent in STVP. This does *not* mean that the principle of tolerance, interpreted strictly, is equivalent to its tolerant interpretation. Indeed, a prima facie

secondly, the closeness constraint proper that if $M \models^{s/t} aI_Pb$, then $|I(Pa) - I(Pb)| < 1$ (not assuming the "only if" part). That way of doing things actually appears preferable to us in general, but we collapse both constraints in 21.1 for the sake of simplicity.

Fig. 21.2 Mixed consequence for vague predicates



counterintuitive consequence of our account is that $Pa_1, a_1 I_P a_2 ... I_P a_n, \forall xy(Px \wedge xI_P y \supset Py) \models_{STVP} Pa_n$. That is, if the tolerance principle is assumed to hold strictly, then the sorites paradox shows its ugly head again. However, the tolerance principle is only tolerantly valid, and cannot be used as a strict premise to derive new consequences. Note that this consequence is as it should be. For the tolerance principle $\forall xy(Px \wedge xI_P y \supset Py)$, interpreted strictly, actually means that if P holds *tolerantly* of x, and x and y are indiscriminable, then P holds *strictly* of y. Viewed in this way, we see that it now is a much stronger principle than when interpreted tolerantly. Only the tolerant interpretation, in our view, captures the adequate pretheoretical meaning attached to the notion of tolerance.

21.3.2 STTT

The hallmark of a truth predicate on our account is the *transparency principle*, according to which a sentence A should be intersubstitutable for $T\langle A \rangle$ in all extensional contexts and in all arguments without change of validity. Our reasons to hold on to transparency are fundamentally to let truth fulfill its expressive function in natural language (see (Field 2008) and (Cobreros et al. 2013) for ampler discussion). In two-valued classical logic, however, and provided the language is sufficiently expressive, the transparency principle leads to the Liar paradox and other related paradoxes, such as the Curry paradox.

Contrary to the tolerance principle in the case of vagueness, which is often viewed with suspicion by supporters of two-valued classical logic, the transparency principle for truth is generally seen as desirable even by supporters of bivalent classical logic. Because of that, a family of responses to the paradoxes of truth consists in typing truth predicates (a move first made by Tarski). This, intuitively, corresponds to one way of limiting the expressiveness of the language: sentences like the Liar or the Curry sentence are not well-formed. An alternative, first explored by (Kripke 1975), consists in changing the logic, without the need for type-distinctions.[6] Kripke's approach thus succeeds in preserving the transparency principle, but it has several limitations. One of those concerns the fact that the resulting logic, K3TT (for K3 with Transparent Truth), is too weak to validate other principles that seemingly ought to result from transparency, such as the T-equivalence $T\langle A\rangle \equiv A$. In this section we show that by adopting ST as our background logic, we can likewise achieve transparency for truth, but without falling prey to the same limitations. As in the case of the sorites paradox for vagueness, the approach diagnoses the Liar and kindred paradoxes as making illegitimate use of the transitivity of logical consequence.

To see this, we proceed to define the system STTT (for ST with Transparent Truth). As our language in what follows, we assume the language of first-order logic without identity and function symbols, augmented with a distinguished predicate T for truth, and with a quote-name forming operator $\langle\rangle$, such that $\langle A\rangle$ is a name for the sentence A. In this language, in particular, we assume that we can formulate self-referential sentences such as the Liar sentence λ, which by definition is the sentence $\neg T\langle\lambda\rangle$, or the Truth-teller sentence τ such that τ is the sentence $T\langle\tau\rangle$, or the Curry sentence κ identical to $T\langle\kappa\rangle \supset A$ (for A a sentence that may take value 0 on all models). Our models for this language are Kripke-Kleene models, namely three-valued models of the same kind used so far, but with the following two constraints:

$a.$ $\langle A\rangle$ always denotes the sentence A

$b.$ A and $T\langle A\rangle$ always have the same truth value (identity of truth) (21.4)

Note that the identity constraint on truth (also called the *fixed point property*) is an essential component toward transparency in our theory, but that identity and transparency are two independent constraints in general.[7] Likewise, closeness as defined in 21.1 is a component toward tolerance in the present theory of vagueness, but closeness and tolerance too are independent principles.[8] In the architecture of our theory therefore, we may say that the identity constraint on truth has the same

[6] Type-free treatments of the Liar purporting to maintain classical logic ought to be mentioned too. See in particular the contextualist accounts of (Parsons 1974) and (Glanzberg 2004), who both argue that the Liar rests on a phenomenon of variable quantifier domain restriction.

[7] We are indebted to an anonymous reviewer for this important clarification. As pointed out by the reviewer, in Kripke's construction the minimal fixed point V for the supervaluation schema satisfies the fixed point property, but not transparency, since as a consequence of the lack of value-functionality in the supervaluation schema, $V(\lambda \vee T\langle\lambda\rangle) = V(\neg T\langle\lambda\rangle \vee T\langle\lambda\rangle) = 1$, but $V(T\langle\lambda\rangle \vee T\langle\lambda\rangle) = 1/2$. Conversely, a transparent theory of truth may fail identity, for example if you start from a Kripke-Kleene model M and generate a new model M' that assigns to each sentence A the pair $<M(A), A>$ as a value, and then simply ignore the second coordinate of its values when defining validity. This sort of model will yield the same logic as the original models, but without ever assigning the same value to any two distinct sentences, so it will exhibit transparency without identity.

[8] (Smith 2008) presents closeness as an explicit weakening of tolerance in his fuzzy approach, and means to endorse closeness without endorsing tolerance. See (Cobreros et al. (Ms.)) however for a more thorough discussion of the status of both principles in relation to Smith's notion of consequence. Conversely, the theory of vagueness presented in (Cobreros et al. 2012b), which

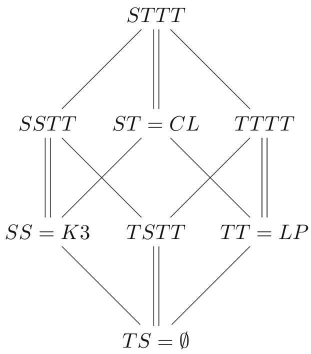


Fig. 21.3 Mixed consequence for transparent truth

priority with regard to transparency as the closeness constraint does with regard to tolerance in the case of vague predicates. That is, identity and closeness are initial model-theoretic postulates governing our special vocabulary, from which we are able to derive transparency and tolerance as general principles governing validity (even though identity and closeness are not meant to be substantially analogous besides this functional level).

That models satisfying (4)-a and (4)-b exist results from Kripke's 1975 fixed-point construction. The main difference with Kripke's approach is that we define logical consequence in terms of strict-to-tolerant consequence, that is, a sentence A follows from a set Γ of formulae provided there is no model where all the formulae of Γ take value 1 and where A takes value 0. Like the Strong Kleene definition of consequence (or indeed the LP one, or the tolerant-to-strict), this notion of consequence supports transparency (see (Ripley 2012) for a proof of this result). One particular consequence of this is the fact that all T -equivalences are STTT-valid. One of the essential benefits of this choice, moreover, which sets it apart from the other schemes, is that if an inference involving a T-free sentence is classically valid, then it remains STTT-valid for any uniform substitution over the full vocabulary (see (Ripley 2012)). Furthermore, the logic is simply better behaved than other three-valued logics in its vicinity, like LPTT, in which transparency and the T-equivalences can be validated, but where the rule of modus ponens is lost. Moreover, with regard to the conditional, STTT satisfies the deduction theorem, which neither K3TT nor LPTT do (compare the situation with the case of vagueness). If we map the extensions of TS, SS, TT and ST that we obtain with the enforcement of transparency, we get a diagram exactly congruent to the one we had for the corresponding extensions with vague predicates, as shown in Fig. 21.3.

As in the previous case, in Fig. 21.3 double lines (read top-down) in the figure indicate (model-theoretic) conservative extensions, and simple lines (top-down) that

involves the notion of classical extension for vague predicates, is one in which tolerance is st-valid, but without involving the notion of closeness in truth values. Tolerance t-holds in all models despite the existence of elements a and b for which aI_Pb holds (strictly or tolerantly), but such that $|I(Pa) - I(Pb)| = 1$ in the two-valued models used.

one logic contains strictly more validities than the other. Note that in contrast to our treatment of vague predicates, the inclusion of a transparent truth predicate does not necessarily add more validities. For instance, TSTT remains an empty consequence relation, just like TS (contrast this with STTT, in which A follows from $T\langle A\rangle$—a schema that would not be valid over plain ST augmented with T and quote-name operators but without the two provisos on names and identity).

Just like STVP, STTT no longer exactly coincides with two-valued classical consequence, despite preserving so many of the features of the latter. First of all, STTT validates sentences that would not be classically valid, on pain of contradiction. One such validity is the Liar sentence λ. The Liar is a valid sentence in STTT because it can only take the value $\frac{1}{2}$ (as in any Kripke fixed point), and therefore cannot take the value 0. Relatedly, STTT departs from standard classical logic in that it is not a transitive consequence relation. Thus, we have that for every sentence A, $A \models_{STTT} \lambda$ and for any sentence B, $\lambda \models_{STTT} B$, but we do not have $A \models_{STTT} B$ for every sentences A and B. Note that, as in the case of the tolerance principle for STVP, the loss of transitivity in STTT precisely explains why counting the Liar as a valid sentence is compatible with blocking the Liar paradox. In particular, we have that $\models_{STTT} T\langle\lambda\rangle \wedge \neg T\langle\lambda\rangle$. That is, we can infer both that the Liar is true and that it is not true. Likewise, we do have that $T\langle\lambda\rangle \wedge \neg T\langle\lambda\rangle \models_{STTT} A$, for any sentence A. If the Liar is true and not true, then anything follows. But we cannot derive $\models_{STTT} A$, that is, we cannot derive that any formula can be accepted tolerantly. The reason is that we illegitimately chained two valid inferences here, but one in which $\lambda \wedge \neg\lambda$ is accepted tolerantly as a conclusion with one where it is used strictly as a premise for further reasoning.

The analogy with our treatment of tolerance in STVP is worth stressing. Remember that in STVP, tolerance (whether as an axiom, or as a rule) was tolerantly but not strictly valid. Hence, it could not be used as a sound premise in an STVP-valid argument. Similarly here, one can accept tolerantly that the Liar is true and not true, but assuming strictly that it is true and not true (or even just one of them) leads to contradiction. Again, note that this consequence is as it should be. For this means that the Liar is not a sentence one can accept or deny strictly. But it remains a sentence that can be accepted tolerantly.

21.4 Nontransitivity and Revenge

In the previous section we have shown how to deal with the sorites paradox and the Liar paradox by means of strict-to-tolerant consequence. In this section, we propose an assessment of our approach. Two main issues need to be considered. The first concerns the scope of permissive consequence, and the question of how high a cost it is to give up transitivity for consequence in the face of the paradoxes. The second concerns whether we can similarly deal with higher-order versions of the paradoxes, namely with strengthened liars and higher-order vagueness.

21.4.1 Nontransitive Consequence

We have seen that both STTT and STVP are nontransitive consequence relations. One objection that could be made to our approach is that transitivity is too intimately tied to the analytic notion of consequence for our approach of the paradoxes to count as satisfactory.

Our answer to this objection is that in the case of STTT as well as in the case of STVP, transitivity is lost only to the extent that some conclusions can be drawn from strict premises that can only be accepted tolerantly. In other words, transitivity is lost only where it would be illegitimate to feed those sentences which can only be accepted tolerantly as strict premises to further arguments. However, transitivity is retained wherever we can ensure that we go from strictly accepted premises to strictly accepted conclusions. This would typically happen where we deal with non-vague predicates in particular (so in large chunks of science, where we can ensure precision). This also happens wherever we deal with sentences that do not involve the truth predicate at all. Arguably, therefore, the loss of transitivity in our system is quite limited: it only affects those inferences that involve special vocabulary, such as the truth predicate, or similarity predicates.[9]

Of course, the question may be asked of how good it is, then, to confer a special status of tolerant validities to sentences such as the tolerance principle, or such as the Liar. Why not simply work with only one notion of assertion, strict assertion, stick to transitive consequence for it, and rest content with the view that the tolerance principle or the Liar are paradoxical sentences, which should simply be excluded from sound reasoning?

This question raises fundamental issues, probably too fundamental to be answered satisfactorily here. However, one methodological answer is that, in so far as identity (see 21.4) is considered a basic postulate for truth on our account, and likewise in so far as the principle of closeness (see 21.1) is taken to govern our intuitions about similarity for vague predicates, going with strict assertion only would simply prevent us from getting any object-language equivalent of these principles as validities (such as the T-equivalences, or the tolerance principle). A more fundamental intuition we have is that going with strict assertion only would put too high standards on assertion. In the case of vagueness, for example, we do not agree that borderline cases of P, because they are neither strictly P nor strictly not P, should command silence on the part of speakers. Rather, we think that borderline cases are cases for which we have equally good reasons to issue judgments either way (see (Wright 1995), (Raffman 2014), (Cobreros et al. 2012b), (van Rooij 2012), (Egré 2011), (Ripley 2013a) for distinct but compatible justifications for this view). In the case of truth, mutatis mutandis, we consider the Liar to be a sentence for which there are inferential reasons for acceptance as well as rejection. Simply refraining to assert anything of

[9] See also (Cobreros et al. 2013) for a discussion of structural motivations for the admission of non-transitive consequence.

the Liar because of that would seem to us to amputate those inference grounds.[10] Of course, the question can be posed again of why it is not good enough to work with only one notion of assertion, namely the tolerant notion. Our answer in this case is that the logic we get is too weak. Consider vagueness again: in the LP version of our approach, the inference from Pa, aI_Pb to Pb is not valid, although the corresponding conditional is. This discrepancy appears to undercut the very motivation for having a tolerance principle.[11]

More specific worries may be expressed still about the failure of transitivity in STVP or STTT. In particular, one way in which the failure of transitivity shows in STTT and STVP concerns the closure of validities under modus ponens. In STTT, since $\neg\lambda$ is tolerantly valid, so is any sentence of the form $\lambda \supset A$, with A an arbitrary sentence. However, we cannot detach A for any such A, as soon as A is a sentence that can be denied strictly. So the Liar is a sentence that gives us as many conditionals as we want, but many of those will be conditionals whose consequent cannot be guaranteed to hold tolerantly (in contrast to the consequent Pb of the non-trivial conditional $Pa \wedge aI_Pb \supset Pb$). Likewise, the set of STVP validities is not closed under modus ponens either. For example, $(Pa \vee \neg Pa) \supset ((aI_Pb \wedge aI_Pc) \supset (Pb \supset Pc))$ is STVP-valid, and so is $Pa \vee \neg Pa$. But $((aI_Pb \wedge aI_Pc) \supset (Pb \supset Pc))$ is not (let Pc take value 0, Pb take value 1, and Pa take value 1/2). Here too, we get the example of a conditional sentence whose validity can only be useful if the antecedent can be asserted strictly. In case the antecedent is only tolerantly assertable (as stipulated here), the sentence's tolerant validity makes the sentence too fragile for consequences to be detached safely.

A further concern one might have toward our non-transitive notion of consequence is how we can know of an arbitrary sentence that it can only be asserted tolerantly. After all, whether a sentence is Liar-like is contingent (see Kripke's Nixon-Dean example). So when we utter a sentence, how can we know that we are not allowed to assert it strictly? How can we know in particular that we are not allowed to reason transitively with a given sentence? Our answer to this question is that, as far as possible, our commitments with regard to assertion and reasoning should go with strict standards. If we know our grounds are safe enough for a strict assertion, then we can use modus ponens transitively provided we know the corresponding conditional itself to be good enough.[12] In other words, what ST-consequence recommends is:

[10] On the comparison between norms of assertion with regard to theories of truth, see especially (Wintein 2012). Chapter 7 of (Wintein 2012), in particular, presents a theory of truth based on the strict-tolerant distinction, but taking a different perspective on Kripke-Kleene models for truth as well as on assertibility proper.

[11] See, again, (Zardini 2008).

[12] Our view on this should be compared to Priest's original view on the status of modus ponens, a rule that is not LP-valid, but that Priest calls a "quasi-validity", still applicable to sentences that are not paradoxical. In our system, modus ponens is a validity, but wherever Priest talks of quasi-validities that are lost in relation to the conditional, we can speak of corresponding classical metainferences that are lost for consequence in the vicinity of paradoxes.

make sure that your conclusions are sufficiently robust in order to start using them as new premises.

Whether we are justified to assert a sentence strictly may not always be easy to ascertain, however. Because of that, we have to agree that our theory does not provide any a priori characterization of those sentences that can be asserted strictly, as opposed to tolerantly only. Upon reflection, however, the problem may be no more nor less pressing than it is when dealing with a transitive consequence relation. Suppose we had elected K3 as our logic. The choice of K3, a transitive consequence relation, would not make the predicament of determining whether a sentence is grounded (hence assertible strictly, or deniable strictly) easier to solve than it is with ST as our logic.[13] As argued by Kripke, any adequate theory (transitive or not) needs to admit an element of "risk" when dealing with truth and paradoxical sentences. But still, one could argue that the need to care about which sentences are assertible or deniable only matters for the soundness of arguments when our logic can rely on a unique mode of assertion, but that it does not matter for validity. In contrast, in ST we need to make sure that a sentence is assertible with the same force throughout in order to chain inferences.

This is indeed the case, but note that even in the setting of a classical and transitive logic, care needs to be taken in order to avoid ambiguity, and ambiguity is always likely to disrupt the validity of an argument. Here, we are talking of content-level ambiguity. Consider the following argument: "Aristotle is a merchant; if Aristotle is a merchant, then Plato is a slave. Hence Plato is a slave". For this inference to be valid, we need to ascertain that the names "Aristotle", "Plato", and the predicates "merchant" and "slave" get the same meaning in each occurrence. Usually, we assume such content-level ambiguities as already filtered out. But here too, that is with regard to reasoning quite generally, there is an element of risk. As a matter of principle, even for a transitive logic, validity holds only if we are certain to have avoided any equivocations. We take it that the problem is no more dramatic once we introduce two modes of assertion. In that case, the ambiguity concerns the mode rather than the content of sentences, but we view it as a virtue of our logic, rather than a defect, that it makes explicit the way in which ambiguities are likely to affect argument-validity quite generally. (Lewis 1982) famously writes: "Logic for ambiguity—who needs it? I reply: pessimists." Our enterprise may be called: logic for assertoric ambiguity, but this is not to endorse pessimism about equivocation, since unlike the logics discussed by Lewis, our logic is explicitly committed to a dual theory of assertion.

[13] We are indebted to an anonymous reviewer for drawing attention to this point.

21.4.2 Revenge Issues: Strengthened Liars and Higher-Order Vagueness

We now turn to the second main objection our account needs to face. The objection concerns the recurrence of paradoxes at higher-orders. Whether for vagueness or for truth, this objection is usually pressed against three-valued accounts quite generally, irrespective of how logical consequence is defined in them.

Consider vagueness first. First-order vagueness is the claim that, between the clear instances of a predicate, and the clear counter-instances, there are borderline cases. Second-order vagueness in particular is the claim that there should also be borderline cases of borderline cases. That is, there should not be a sharp cutoff between clear cases of P and borderline cases of P. It may appear, however, that we are committed to such a sharp cutoff by accepting, in our models, the existence of at least two individuals d and d' in a model such that they are P-similar, and yet such that $I(P)(d) = 1$ and $I(P)(d') = 1/2$. Another way to phrase the problem is the following. Introduce an operator D for "determinately", such that $I(DA) = 1$ if $I(A) = 1$ and $I(DA) = 0$ if $I(A) < 1$. Clearly, $M \models^s DA$ iff $M \models^t DA$. Now, take any individual a such that $I(Pa) = 1/2$, that is, a is borderline P. Necessarily, $M \models^{s,t} D(\neg DPa \wedge \neg D\neg Pa)$, that is: a has to be a clear borderline case of P according to that definition.

Similarly, in the case of truth, although the Liar sentence can take on the value $1/2$ without contradiction, and without threatening transparency, a strengthened version of the Liar brings contradiction back in if we accept to enrich our vocabulary. Again, let D be the determinateness operator such that DA gets value 1 if A gets value 1 in the model, and gets value 0 otherwise. Let σ be the sentence such that σ is equivalent to $\neg DT\langle\sigma\rangle$. Thus, σ says of itself that it is not determinately true. We cannot assign σ a coherent truth value in the model while maintaining transparency anymore, if indeed sentences are allowed to take exactly one of the three truth values at our disposal.

When it comes to revenge issues, theories of vagueness as well as truth are usually faced with a dilemma. One horn of this dilemma is to limit the expressive power of the theory, and to deem unnecessary or illegitimate the introduction of such definiteness operators. The other horn is to consider that expressiveness should not be limited, but that such operators should be treated with particular care. To conclude this paper, we wish to explain in what sense we think our theory is compatible with both horns of this dilemma. Nevertheless, there is a sense in which, by referring the strict and tolerant distinction to assertion, rather than truth, our theory fits maybe more naturally with the idea of preventing the expression of revenge.

The issue of expressive limitation invites a more careful examination of determinateness operators in relation to our framework. One important observation to make about our whole approach is that determinateness operators are not part of the content of the sentences we are interested in, although something like determinacy operators is implicitly at play in the strict-tolerant distinction upon which our theory is built. Indeed, consider an atomic sentence like Pa, meaning that "a is rich". If Pa

holds strictly in our model, then this could be taken to mean that a is determinately rich. Dually, for Pa to hold tolerantly could be taken to mean that a is not determinately not rich. As a matter of fact, instead of introducing strict and tolerant levels for the truth values of our sentences, we could decide to work with a single notion of assertion, but to translate the strict and tolerant metalanguage distinctions into appropriate modal sentences of our object-language, enriched with determinateness operators (see (Kooi and Tamminga 2013) for an exact statement of such a modal translation for basic LP and K3 sentences). For example, to say that the tolerance principle is tolerantly valid, in modal terms, would turn out to be equivalent to the observation that the following "gap principle" holds classically in our models:[14]

$$\forall xy(DPx \wedge xI_P y \supset \neg D\neg Py) \tag{21.5}$$

Although such a translation is available and can be used to embed our treatment of vagueness in modal terms, we think that such an interpretation would likely distort the philosophical motivation of our approach. The main reason is that for us, strict and tolerant are primarily modes of assertion or acceptance; they qualify the force rather than the content of an assertion. Because of that, to assert a sentence such as Pa strictly is not analytically equivalent to the assertion that a is determinately P. Rather, asserting strictly is primarily tied to inferential and coherence commitments (such as the impossibility to deny even tolerantly). Asserting tolerantly, on the other hand, is also to take commitments (such as refusing to deny strictly). Because our approach relies primarily on such speech act distinctions, we thus believe the need to deal with determinateness operators in the object-language is less pressing than for other theories in which semantic values are primarily seen as ways of encoding the relation of a sentence with the world (see (Cook 2009) for such a conception).

Besides, the modal translation does not straightforwardly extend to sentences involving self-reference and the truth predicate. Kooi and Tamminga show how every propositional sentence of the basic language of LP or K3 can be translated into a modal sentence of S5 in a way that preserves argument-validity, but they do not provide a similar translation for the extended language with truth predicates. In LPTT, for example, we know that the sentence $T\langle\lambda\rangle \wedge \neg T\langle\lambda\rangle$ is valid. But if we apply the same translation manual to sentences like the Liar, with no special proviso on the truth predicate, we would get $\neg D\neg T\langle\lambda\rangle \wedge \neg DT\langle\lambda\rangle$ as our translation, and the grounds on which such a sentence should come out modally valid remain to be worked out (under a possible worlds semantics for the D operator).

This does not mean that issues of determinateness should be ignored. Consider the problem of higher-order vagueness. We can still accommodate determinateness operators in the object-language of our theory. But importantly, we do not think that "a is determinately P" should then necessarily have the truth conditions given above. In the case of vague predicates, we could accept, in particular, that the D operator does not necessarily map three-valued sentences to 0 or 1, but that it can

[14] See (Wright 1992), (Fara 2003), (Cobreros 2011) on gap principles, and (Egré 2011) and (van Rooij 2012) on the link between tolerance in the strict-tolerant framework and gap principles.

map sentences to 1/2, depending on the case (so that one could strictly assert that someone is bald, and still say something different with the strict assertion that that person is determinately bald). Whether we can accommodate indefinitely iterated borderline cases is an issue that goes beyond the scope of this paper, but our main point, once more, is that the semantics of determinateness is a matter distinct from strict assertion proper.

The situation is more thorny in the case of the strengthened Liar. We have to agree that, from a metatheoretical point of view, the strengthened Liar remains unescapable as soon as the relevant expressive means are available, just as, in recursion theory, any attempt to give a complete specification for the set of recursive functions, not involving partial functions, is threatened by the diagonalization method (Rogers 1987).[15] As (Priest 1984), (Cook 2009), and (Schlenker 2010) have argued, the strengthened Liar may in fact be considered an argument for the idea that truth values are indefinitely extensible, and that working with only three truth values sets an artificial bound on this phenomenon of indefinite extensibility. According to Cook and Schlenker, in particular, given any set S of truth values, the problem will indeed recur as soon as we have a sentence ρ equivalent to "ρ has a truth value in S other than true".

However, we have to emphasize that, in our theory, the natural way to express the sentence saying of itself that it is "other than true" is the standard Liar sentence. The reason is that we do not see the predicate "True" as tied to the value 1. Rather, "True" on our account is a predicate whose function is primarily inferential (as reflected by the identity constraint). In principle, however, we can still build a sentence that says of itself: "I am not strictly assertible", which one may formalize in terms of the sentence σ given above. Once we let sentences such as σ in, what are we to do with them? One possible line of response is to assent to the view of the indefinite extensibility of truth values, but to maintain the principled division between two modes of assertion. To do this, we may use a construction proposed by (Priest 1984).[16] To deal with σ, enlarge the space of truth-values to the power set of $\{0, \frac{1}{2}, 1\}$ (minus the empty set), and repeat the construction at higher levels. Now, consider what happens if σ gets the value 1: then it has to get value 0, and conversely. If it gets the value $\frac{1}{2}$, then it has to get the value 1, and also 0. By this reasoning, it seems the possible values for σ, upon this extended set, are $\{1, 0\}$ and $\{1, \frac{1}{2}, 0\}$. Now, let us say that, relative to this set, a sentence A is tolerantly assertible if the values it gets all contain 1 or $\frac{1}{2}$, tolerantly deniable if the values all contain 0 or $\frac{1}{2}$, strictly assertible if its value is $\{1\}$, and strictly deniable if its value is $\{0\}$. Seen in that way, the sentence "I am not strictly assertible" is therefore tolerantly assertible and deniable, and neither strictly assertible nor strictly deniable. This appears to be

[15] Rogers' emphasis on the use of partial, as opposed to total, functions as a way of blocking diagonalization arguments bears some analogy with the idea of limiting the expressiveness of our language to block the strengthened Liar.

[16] See (Ripley 2013a) for an application of this strategy to deal with a particular version of higher-order vagueness.

a desirable outcome, since by the valuation chosen, σ is not strictly assertible. But because this is what σ says, it appears we should be able to assert it in some sense, which we can do tolerantly. Similarly, consider a sentence ρ saying of itself that it is not tolerantly assertible (hence strictly deniable). This sentence is in fact tolerantly assertible on the valuation chosen. Hence it appears we should be able to deny it in some sense, which we can do tolerantly. Finally, take a sentence ν saying of itself that it is neither strictly assertible nor strictly deniable. If it gets value 1, it has to get value 0, and if it gets value $\frac{1}{2}$, it has to get value 1. However, it can get value 0 without contradiction. So whichever value it takes, it ought to get the value 0. Hence the sentence is strictly deniable. But since the sentence denies this, it seems that one should be able to deny it in some sense, which we can.

Based on this sample of examples, the valuation chosen appears to do justice to intuitions about acceptance and denial based on the inferences we can perform, and in a way that is faithful to inferential practice. Also, we can see that this is a way of replicating the trichotomy between strictly assertible, strictly deniable and tolerantly assertible one level up. Of course, in doing so, we are building a hierarchy for assertibility that parallels the hierarchy of truth values. But the important point is that, for any new predicate that one might introduce in the language to build a new extended Liar, one can come up with a way of understanding strict and tolerant assertion that will make the sentence neither strictly deniable, nor strictly assertible, but tolerantly both. The view of revenge as imposing an indefinite extension of truth-values is therefore compatible with the basic architecture of our theory, that is with the distinction between the mode of assertion (strict or tolerant) and the predicates intended to reflect those properties of assertibility in the object-language.

21.5 Conclusion

Our account of the paradoxes in terms of permissive consequence presents several advantages over other three-valued accounts in its vicinity. Firstly, with regard to other three-valued accounts based either on paracomplete or paraconsistent solutions, it allows us to maintain a simple logic, with a simple conditional, an ingredient that is notoriously missing from standard three-valued theories of either truth or vagueness. This feature, as we have emphasized, is obtained essentially because of the duality between strict and tolerant interpretations in our system, a duality that is missing from Kleene's logic as well as from LP.

Secondly, the framework accounts both for the tolerance of vague predicates, and for truth obeying the T-equivalence (on top of transparency) without paradox. From a philosophical point of view, however, it is particularly important to point out that tolerance and the T-equivalences are only tolerant validities. They are not strict validities, on pain of contradiction. In this, our framework, as the name indicates, is permissive indeed, but it comes with a caveat that inferences based on those principles in conditional form are fragile, because they fundamentally require the antecedent to be accepted strictly in order to go through safely.

The third feature of our account is that, because permissive consequence involves a shift of standard from premises to conclusions, it has to give up on the transitivity of logical consequence. Opponents to our account will likely consider that by validating tolerance or the T-schema without restriction, we have traded a reliable notion of consequence for principles that are fragile and of limited use. But this conclusion would be unfair. By its emphasis on two modes of acceptance, strict and tolerant, our account of consequence simply recognizes that chaining inferences is not an innocent business. As soon as we make room for reasoning with either vague predicates or self-referential sentences, logic, on our view, and logical consequence with it, needs to incorporate more generality, and more complexity with it.

Acknowledgments We are grateful to two anonymous referees for detailed comments, and to the editors for their assistance in the preparation of this paper. We also thank audiences in Amsterdam, Barcelona and Paris. We thank the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n 229 441-CCC, the ANR project BTAFDOC ("Beyond Truth and Falsity: Degrees of Confidence"), and the program "Non-Transitive Logics" (Ministerio de Economía y Competitividad, Government of Spain, FFI2013-46451-P). We also thank grants ANR-10-LABX-0087 IEC and ANR-10-IDEX-0001-02 PSL*.

References

Bennett, B. (1998). Modal semantics for knowledge bases dealing with vague concepts. In A. G. Cohn, L. Schubert, & S. Shapiro (Eds.), *Principles of Knowledge Representation and Reasoning: Proceedings of the 6th International Conference (KR-98), Morgan Kaufman, 1998*.

Cobreros, P. (2011). Supervaluationism and Fara's argument concerning higher-order vagueness. In P. Egré & N. Klinedinst (eds.), *Vagueness and language use* (pp. 207–221) Palgrave Macmillan.

Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2012a). Tolerance and mixed consequence in the s'valuationist setting. *Studia Logica, 100*(4):855–877.

Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2012b). Tolerant, classical, strict. *Journal of Philosophical Logic, 41*(2), 347–385.

Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (Ms). Tolerance and Degrees of Truth. Manuscript, in progress (of a paper presented at Cerisy 2011 and ESSLLI 2012).

Cobreros, P., Egré, P., Ripley, D., & van Rooij, R. (2013). Reaching transparent truth. *Mind. 122*(488):841–866.

Colyvan, M. (2009). Vagueness and truth. In H. Dyke (ed.), *From truth to reality: New essays in logic and metaphysics* (pp. 29–40). Routledge.

Cook, R. (2009). What is a truth value and how many are there? *Studia Logica, 92*(2), 183–201.

Egré, P. (2011). Perceptual ambiguity and the sorites. In R. Nouwen, R. van Rooij, U. Sauerland, & H.-C. Schmitz (eds.), *Vagueness in communication* (pp. 64–90). Springer.

Fara, D. (2003). Gap principles, penumbral consequence, and infinitely higher-order vagueness. *Liars and heaps: New essays on paradox* (pp. 195–222) (Originally published under the name "Delia Graff").

Field, H. (2003). Semantic paradoxes and vagueness paradoxes. In J. Beall (ed.), *Liars and heaps* (pp. 262–311). Oxford: Oxford University Press.

Field, H. (2008). *Saving truth from paradox*. Oxford: Oxford University Press.

Frankowski, S. (2004). Formalization of a plausible inference. *Bulletin of the Section of Logic, 33*(1), 41–52.

Glanzberg, M. (2004). A contextual-hierarchical approach to truth and the liar paradox. *Journal of Philosophical Logic, 33*(1), 27–88.
Hyde, D. (1997). From heaps and gaps to heaps of gluts. *Mind, 106,*641–660.
Kleene, S. (1952). *Introduction to metamathematics*. D. Van Nostrand Company Inc.
Kooi, B., & Tamminga, A. (2013). Three-valued logics in modal logic. *Studia Logica*, *101*(5), 1061–1072.
Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy, 72*(19), 690–716.
Lewis, D. (1982). Logic for equivocators. *Noûs, 16*(3)*,* 431–441.
Malinowski, G. (1990). Q-consequence operation. *Reports on mathematical logic, 24,* 49–54.
McGee, V. (1991). *Truth, vagueness, and paradox: An essay on the logic of truth*. Hackett Pub Co Inc.
Nait-Abdallah, A. (1995). *The logic of partial information*. Springer.
Parsons, C. (1974). The liar paradox. *Journal of Philosophical Logic, 3*(4), 381–412.
Priest, G. (1979). Logic of paradox. *Journal of Philosophical Logic, 8,* 219–241.
Priest, G. (1984). Hyper-contradictions. *Logique et Analyse, 27*(107), 237–243.
Priest, G. (2010). Inclosures, vagueness, and self-reference. *Notre Dame Journal of Formal Logic, 51*(1), 69–84.
Raffman, D. (2014). *Unruly words: A study of vague language*. Oxford University Press.
Ripley, D. (2012). Conservatively extending classical logic with a transparent truth predicate. *Review of Symbolic Logic 5*(2):354–378.
Ripley, D. (2013a). Sorting out the sorites. In F. Berto, E. Mares, & K. Tanaka, (eds.), *Paraconsistency: Logic and applications*. Springer 329–348.
Ripley, D. (2013b). Paradoxes and failures of cut. *Australasian Journal of Philosophy* 91(1): 139–164.
Rogers, H. (1987). *Theory of recursive functions and effective computability*. MIT.
Rumfitt, I. (2000). "Yes and no". *Mind, 109*(436), 781–823.
Schlenker, P. (2010). Super liars. *The Review of Symbolic Logic, 1*(1), 1–41.
Smith, N. J. J. (2008). *Vagueness and degrees of truth*. Oxford: Oxford University Press.
Tappenden, J. (1993). The liar and sorites paradoxes: Toward a unified treatment. *The Journal of Philosophy, 90*(11), 551–577.
van Rooij, R. (2012). Vagueness, tolerance, and non-transitive entailment. In P. Cintula, C. Fermueller, L. Godo, & P. Hàjek (Eds), *Understanding vagueness—Logical, philosophical and linguistic perspectives* (pp. 205–221). King's College Publications.
Williamson, T. (1994). *Vagueness*. London: Routledge.
Wintein, S. (2012). *Playing with truth*. PhD Thesis, Tilburg University.
Wright, C. (1976). Language-mastery and the sorites paradox. In G. Evans, & J. McDowell (Eds.), *Truth and meaning: Essays in Semantics*, pp. 223–247.
Wright, C. (1992). Is higher order vagueness coherent? *Analysis, 52*(3)*,* 129–139.
Wright, C. (1995). The epistemic conception of vagueness. *The Southern journal of philosophy, 33*(S1), 133–160.
Zardini, E. (2008). A model of tolerance. *Studia Logica, 90,* 337–368.